



Feldabgrenzung und internationale Kollaborationen: Praktische Begleitforschung zum „Pilotvorhaben zur Identifizierung internationaler Kooperationsmöglichkeiten“ des BMBF – Abschlussbericht

Kompetenzzentrum Bibliometrie

Institut für Forschungsinformation und Qualitätssicherung (iFQ)

Schützenstraße 6a
10117 Berlin

Telefon 030-2064 177-0

Fax 030-2064 177-99

E-Mail info@forschungsinfo.de

Internet www.forschungsinfo.de | www.research-information.de

April 2013

Inhalt

1	Projektkontext, Forschungsbedarf und Fragestellungen	2
2	Zusammenfassung der Ergebnisse	4
2.1	Feldabgrenzung	4
2.2	Kollaborationen	5
2.3	Skalenunabhängige Indikatoren.....	6
2.4	Wissenschaftsdynamik	8
3	Literatur	11
4	Anhänge	14

1 Projektkontext, Forschungsbedarf und Fragestellungen

Im Februar 2012 hat das iFQ unter Mitarbeit des IWT dem Internationalen Büro des BMBF den Prototypen eines bibliometrischen Monitoring-Instruments für internationale Kooperationslandschaften überreicht.¹ Dieser soll zwei wesentlichen Anforderungen gerecht werden: Erstens, er muss beliebige Felder (bestehende Disziplinen und Fächer sowie junge und hochdynamische Forschungsgebiete) abbilden können. Zweitens muss er Indikatoren für verschiedene Leistungsdimensionen und Untersuchungsebenen bereitstellen. Zur bibliografischen Abgrenzung von Feldern wurde ein Protokoll entwickelt, mit dem Felder unter Verwendung von Referenzinformationen (z.B. im Feld verwendete Stichworte, relevante Zeitschriften und Konferenzen oder bekannte Schlüsselpersonen) über zwingend in Titel, Abstract oder Schlagwörtern vorkommenden Suchterme abgegrenzt werden. Zur Leistungsbewertung wurden die folgenden Indikatoren implementiert: In der Dimension *Produktivität* die Anzahl der Publikationen, deren bewertete Wachstumsrate (Sharpe Ratio) und die Spezialisierung; in der Dimension *Kooperation* der Anteil internationaler Kooperationen und die Wichtigkeit eines Zielobjekts als Kooperationspartner (Kooperationspräferenz); in der Dimension *Resonanz* der Anteil unzitierter Publikationen, die mittleren feld- oder zeitschriftnormalisierten Zitatraten und der Beitrag zu den 10% der meistzitierten Publikationen eines Feldes.

Nach Überreichung des Monitoring-Instruments wurden die Abfragen zur Erstellung einer bibliometrischen Analyse der Region Lateinamerika und Karibik genutzt. Die zugehörige Studie befindet sich derzeit in der Abschlussphase.

Schon zu Beginn des Vorhabens wurde deutlich, dass ein solches Projekt grundlegendere Fragen aufwirft, die im Rahmen eines Pilotprojektes nicht bearbeitet werden können. Bei den implementierten Verfahren und Metriken handelt es sich um solche, die den aktuellen Stand der Forschung umsetzen. Teilweise ist die Methodendiskussion jedoch noch nicht abgeschlossen, Weiterentwicklungen sind verfügbar oder nachgefragt. Aus diesem Grund bewilligte der Beirat des Kompetenzzentrums Bibliometrie (KB) im März 2011 das hier adressierte Projekt zur praktischen Begleitforschung. Dieses sollte den beiden grundlegenden Fragen nachgehen:

- Mit welchen Verfahren und Methoden lassen sich komplexe oder emergente Themengebiete bibliographisch abgrenzen?
- Wodurch zeichnen sich geeignete internationale Kooperationspartner auf verschiedenen Ebenen – Land, Institution und Person – aus? Welche Indikatoren können zur Identifizierung interessanter Kooperationspartner herangezogen werden?

Aufgabe war es, ein Verfahren zu entwickeln und zu testen, das Felder unter Verwendung semantischer Verfahren abgrenzt. Darüber hinaus sollten neue Verfahren zur Charakterisierung, Identifizierung und Bewertung potentieller Kollaborationspartner entwickelt und getestet werden. Unter Weiterentwicklung bestehender Verfahren sollten neue Methoden für Feldabgrenzung und Kollaborationsbewertung entwickelt und erprobt werden, um sie für das Monitoring-Instrument nutzbar zu machen.

¹ iFQ, „Pilotvorhaben zur Identifizierung internationaler Kooperationsmöglichkeiten“, Abschlussbericht an das BMBF, 27. Februar 2012.

Zur Bearbeitung der Forschungsfragen erfolgte eine Aufteilung in vier Module:

1. Feldabgrenzung: Entwicklung und Weiterentwicklung von Methoden zur Feldabgrenzung, basierend auf einem Natural Language Processing (NLP)-Verfahren zur Klassifizierung von Publikationen.
2. Kollaborationen: Entwicklung und Weiterentwicklung von Methoden zur Messung des Ertrags von Kooperationen, basierend auf netzwerkanalytischen Verfahren.
3. Skalenunabhängige Indikatoren: Analyse der Skalenfreiheit des Wissenschaftssystems und Entwicklung und Weiterentwicklung von Indikatoren außerhalb der momentbasierten Statistik.
4. Wissenschaftsdynamik: Entwicklung und Weiterentwicklung von Verfahren, die die zeitliche Entwicklung von Feldern und den darin befindlichen Akteuren abbilden oder quantifizieren.

Zur Bearbeitung des ersten Moduls wurde ein Unterforschungsauftrag an das Institute for Science Networking (ISN) an der Carl von Ossietzky Universität Oldenburg vergeben. Die restlichen Module wurden vom iFQ bearbeitet.

Zu jedem der vier Module ist eine Publikation erarbeitet worden. Im Folgenden werden der Forschungsstand, die Forschungsfragen, die Hauptergebnisse und die Implikationen für das Monitoring-Instrument kurz zusammengefasst. Für Details verweisen wir auf die Publikationen selbst. Zwei der Publikationen sind bereits fertiggestellt und finden sich im Anhang. Die übrigen werden nach Fertigstellung den KB-Partnern zeitnah übermittelt.

2 Zusammenfassung der Ergebnisse

2.1 Feldabgrenzung

Stand der Forschung

Multidisziplinäre und dynamische Felder lassen sich mit Quellenklassifikationssystemen, wie sie in *WoS* und *Scopus* verwendet werden, nicht abgrenzen, da diese relativ trennscharf, zu hoch aggregiert und zu starr sind. Dynamische und emergente Felder müssen daher auf Artikelebene abgegrenzt werden. Die einfachste Möglichkeit ist die Verwendung existierender Artikelklassifikationssysteme wie Physics and Astronomy Classification Scheme (PACS) oder Mathematics Subject Classification (MSC), sofern diese vorhanden sind und eine ausreichende Abdeckung aufweisen (Radicchi & Castellano, 2011; Smolinsky & Lercher, 2012). Zusätzlich kann eine Konkordanz zwischen einer Datenbank mit Artikelklassifikationssystem und einer Zitationsdatenbank hergestellt werden (Strotmann, Zhao, & Bubela, 2009; Strotmann & Zhao, 2010). Der klassische, rein lexikalische Ansatz besteht darin, Suchterme zu definieren und zu kombinieren, die zwingend in Titel, Abstract oder Keywords vorkommen (Maghrebi, Abbasi, Amiri, Monsefi, & Harati, 2011; Noyens u. a., 2003). Solche Methoden tendieren allerdings dazu, zu viele falsch positive Publikationen zu erfassen. Zitatbasierte Verfahren tendieren hingegen dazu, zu viele falsch negative Publikationen zu erfassen. Hybride Verfahren sollen die jeweiligen Nachteile aufheben (Wolfgang Glänzel & Thijs, 2011). Dazu kann entweder mit einem präzisen „Seed“ (einer initialen, lexikalisch bestimmten Publikationsmenge) begonnen werden, dem dann Publikationen hinzugefügt werden, die die meistzitierten Publikationen des Seed zitieren (Zitt & Bassecouard, o. J.). Alternativ wird mit einem unpräzisen Seed begonnen, aus dem feldtypische Textfragmente extrahiert werden (Mogoutov & Kahane, 2007). Dadurch und mittels eines parallelen Zitatclustering von Quellen werden Teilfelder abgegrenzt.

Forschungsfragen

Um multidisziplinäre und dynamische Felder auf Artikelebene abgrenzen zu können, wurde die klassische rein lexikalische Feldabgrenzung um ein Verfahren erweitert, das auf maschinellem Lernen basiert. Solche Methoden sind bislang nicht zur Feldabgrenzung verwendet worden. Das Forschungsprojekt soll folgende resultierende Fragen beantworten: Lässt sich die lexikalische Feldabgrenzung damit zumindest teilweise automatisieren und welchen Beitrag kann das überwachte maschinelle Lernen leisten?

Ergebnisse

Das in der Pilotstudie intellektuell abgegrenzte, interdisziplinäre, emergente Feld „Biotechnologie 2020+“ war Gegenstand der maschinellen Abgrenzung. Die Maschine (Classifier) wurde mit 61 Volltexten trainiert (vom Auftraggeber dem Zielfeld zugeordnete Referenzpublikationen). Es wurde keine Negativmenge verwendet. Es wurden 985 Textfragmente identifiziert, die jeweils in mehr als zwei Referenzpublikationen vorkommen. Mit diesem „Fingerabdruck“ wurden 3,8 Millionen Publikationen verglichen, aus denen unter Verwendung eines Grenzwerts das Zielfeld extrahiert wurde. Dabei hat sich als zielführend erwiesen, in den Fingerabdruck Akronyme aufzunehmen, die in definitiv zum Feld gehörigen Publikationen vorkommen.

Hinsichtlich des Vorhandenseins von Kerndokumenten mit teilweise überlappenden Referenzlisten ist die Abgrenzungslösung weniger kohärent als die in der Pilotstudie erarbeitete intellektuelle Lösung. Die Annahme, dass die intellektuelle und automatische Feldabgrenzung die gleichen Publikationen abgrenzen, hat sich nicht bestätigt. Die Verfahren produzieren überschneidungsfreie Lösungen. Dennoch werden beide Lösungen von Experten etwa gleich gut bewertet (Präzision \approx 50%). Das deutet darauf hin, dass das Feld noch zu jung ist, um kanonisches Wissen generiert zu haben und sehr fließende Grenzen hat.

Die schlechte Antwortrate (7%) ist darauf zurückgeführt worden, dass beim Training Volltexte verwendet worden sind, bei der Suche jedoch Abstracts. Zukünftig sollte auch beim Training mit Abstracts gearbeitet werden.

Implikationen für das Monitoring-Instrument

Die Frage, ob sich eine intellektuelle Feldabgrenzung durch die Verwendung von maschinellem Lernen zumindest teilweise automatisieren lässt, kann noch nicht abschließend beantwortet werden. Klar ist, dass beim Training größte Sorgfalt auf die Auswahl von Positiv- und Negativmengen gelegt werden muss. Beispielsweise reicht es nicht aus, allgemeinen Sprachgebrauch als Negativmenge zu verwenden. Je mehr intellektuelle Kontrolle vermieden werden soll, desto mehr müssen die Dokumente der Negativmenge derart gewählt werden, dass sie das Zielfeld eng abgrenzen aber dennoch außerhalb liegen. Im derzeitigen Entwicklungsstadium des Abgrenzungsprozesses ist eine manuelle Durchsicht der Textfragmente unerlässlich.

Als Teil eines hybriden Feldabgrenzungsansatzes kann das maschinelle Lernen bereits Anwendung finden, beispielsweise zur Identifikation von Textfragmenten. Es empfehlen sich Weiterentwicklungen wie die Implementierung von Routinen zur Reduzierung von Worten auf Wortstämme (Stemming) und die Verwendung von TF/IDF-Verfahren zur Selektion von Textfragmenten.

2.2 Kollaborationen

Stand der Forschung

Studien haben gezeigt, dass Kollaboration die Anzahl der Zitationen der kopublizierten Artikel erhöht (Chinchilla-Rodríguez, Vargas-Quesada, Hassan-Montero, González-Molina, & Moya-Anegón, 2010; W. Glänzel, 2001, 2002; Hsu & Huang, 2011; J. S. Katz & Hicks, 1997; Lewinson & Cunningham, 1991; Narin, Stevens, & Whitlow, 1991; Persson, Glänzel, & Danell, 2004). Um zu analysieren, welchen Erfolg eine Kollaboration hat, wurde die „Relative Citation Eminence“ (RCE) vorgeschlagen (W. Glänzel & Schubert, 2001). Die RCE bewertet Kollaborationen jedoch symmetrisch und ermöglicht keine Einschätzung, welches Land stärker oder nicht von einer Kollaboration profitiert. Ein asymmetrisch verteilter Kollaborationsprofit ist jedoch zu erwarten, da sich Länder in ihrer wissenschaftlichen Zusammenarbeit hinsichtlich ihrer Größe, Zitationskultur, Produktivität und geografischen Nähe unterscheiden (Ding, Foo, & Chowdhury, 1998; W. Glänzel, Schubert, & Czerwon, 1999; W. Glänzel, 2001; J. S. Katz & Hicks, 1997; Luukkonen, Tijssen, Persson, & Sivertsen, 1993; Pao, 1981; Singh, 2005; Zhao & Guan, 2011).

Um den asymmetrischen Zitationsertrag bei Kollaborationen zu messen, wurden das „Citation Rate Increment from the Collaborator“ (CRIC), der „Domestic Citation Rate Comparison when Collaborating“ (DCRCC), die „Domestic Impact Rate Increment when Collaborating“ (DIRIC), der „International Collaboration Gain in Normalized Citation“ (ICGNC) (Lancho-Barrantes, Guerrero-Bote, & de Moya-Anegón, 2012), sowie der „International Collaboration Gain in Normalized Citation“ (ICGNC) (Guerrero Bote, Olmeda-Gómez, & de Moya-Anegón, 2013) entwickelt. Diese Indikatoren weisen jedoch einige methodologische Probleme auf. CRIC, DCRCC, und DIRIC sind nicht feldnormalisiert und verzerren deshalb die Ergebnisse in Richtung der hochzitierten Fächer. Weiterhin werden Koautorschaften mit mehr als zwei Ländern nicht gesondert berücksichtigt. Der Fokus auf reine Zitatsummen im Falle von CRIC und DCRCC ist gegenüber Extremwerten anfällig. ICGNC ist feldnormalisiert, jedoch kann die Konstruktion als Differenz statt als Bruch angefochten werden. Schließlich vergleichen alle Indikatoren Fälle von Kollaborationen mit Fällen ohne Kollaboration.

Forschungsfragen

Kooperationen werden oftmals politisch initiiert. Dabei interessiert die Frage, ob alle Partner von Kooperation profitieren – ob manche Kooperationen also eine Art „wissenschaftliche Entwicklungshilfe“ darstellen. Es werden Indikatoren entwickelt, um den „Collaboration Benefit“ asymmetrisch zu beziffern. Da bei wird der Ertrag als Verbesserung des Impact operationalisiert (steigt der zitationsbasierte Einfluss, so resultiert das in einer höheren Bewertung).

Ergebnisse

Es wurde der Indikator „Collaboration Benefit“ (CB) vorgestellt, der feldspezifische Zitatraten aus Kollaborationen mit einem Zielland ins Verhältnis zu den für das Land üblichen Zitatraten in Kooperation entstandener Publikationen setzt. Da eine Berechnung auf Publikationsebene erfolgt, ist eine Aggregation sowohl über Disziplinen als auch über Länder möglich. Der Indikator wurde in abgewandelter Form auch für den internationalen und heimischen Kooperationsertrag berechnet. Bei der Anwendung des Indikators auf die bibliometrische Datenbank des KB wurde deutlich, dass er aufgrund seiner Konstruktion anfällig gegenüber hochzitierten Publikationen ist (wie alle bestehenden mittelwertbasierten Indikatoren).

Demgegenüber wurde ein zweiter Indikator „Citation Share“ (CS) von internationalen und heimischen Zitaten gestellt, der aufgrund seiner Konstruktion als Anteil von Zitation nicht durch hochzitierte Publikationen verzerrt wird. Dieser Indikator gibt damit eine bessere Vorstellung von der typischen Zitationsverteilung der Kollaboration. Eine Korrelationsanalyse der fünf Faktoren auf Basis der KB-Datenbank zeigte, dass die Indikatoren teilweise untereinander stark korrelieren (CB und $CB^{international}$ $R=0,99$; $CS^{international}$ und CS^{inland} $R=-0,77$), teilweise jedoch kaum wechselseitig erklärbar sind (bspw. CB und $CS^{international}$ $R=0,03$). Die Korrelationen implizieren, dass beide Indikatoren von verschiedenen Faktoren (hochzitierte Publikationen und Inlands-Bias) beeinflusst sind. Für die Analyse von Kollaborationen kann durch die parallele Analyse beider Indikatoren eine von hochzitierten Publikationen nicht beeinflusste Bewertung des internationalen Anteils berücksichtigt werden. Zudem ermöglichen die Indikatoren die differenzierte Analyse der nationalen und internationalen Anteile der durch Kooperation mehr oder weniger erreichten Zitationen.

Implikationen für das Monitoring-Instrument

Die Indikatoren wurden auf Ebene der einzelnen Publikation entwickelt, sind demzufolge grundsätzlich auf mehreren Aggregationsstufen berechenbar. Die derzeitige Umsetzung ermöglicht eine Berechnung der Indikatoren auf Länderebene. Die Anwendung auf tieferen Aggregationsniveaus (bspw. Universitäten innerhalb Deutschlands) ist möglich, erfordert jedoch eine vorherige Zuordnung der Publikationen zu den Universitäten. Zudem ist vor einer Berechnung festzulegen, ob mit „internationalen Zitationen“ weiterhin Zitationen anderer Länder gemeint sind, oder ob die Bezugsgröße mit der Objektebene übereinstimmt (dann sind bspw. auf Universitätsebene alle Zitationen, die nicht aus der eigenen Universität stammen gemeint). Beide Berechnungen sind je nach Erkenntnisinteresse denkbar. Die Indikatoren sind aggregierbar konstruiert und können demzufolge fächer- oder länderübergreifend verglichen werden.

2.3 Skalenunabhängige Indikatoren

Stand der Forschung

Im Monitoring-Instrument – wie in der Bibliometrie insgesamt – werden standardmäßig Indikatoren der parametrischen Statistik verwendet (Lundberg, 2007; Schubert & Braun, 1986; Waltman, van Eck, van Leeuwen, Visser, & van Raan, 2011). Für derartige Indikatoren ist zur Normalisierung eine Skalenkorrektur vorgeschlagen worden, bei der der Erwartungswert größenabhängig ist (J. S. Katz, 2000). In den letzten Jahren sind außerdem nichtparametrische Indikatoren diskutiert und verwendet worden, die den Beitrag zu den oberen Perzentil der meistzitierten

Publikationen beziffert (Bornmann, 2013; Tijssen, Visser, & Van Leeuwen, 2002). Wir nennen diese drei Indikatoren plus den Anteil zitierter Publikationen das „klassische Quadrupel“.

Empirische Zitationsverteilungen sind nicht normalverteilt sondern weisen eine deutliche Schiefe auf (gering zitierte Publikationen bilden den größten Anteil, während vielzitierte Publikationen seltener auftreten). Da Mittelwerte sehr sensitiv für Extremwerte und damit für lediglich einen kleinen Teil der zu bewertenden Publikationen sind, bilden mittelwertbasierte Indikatoren die Fähigkeit eines Objekts ab, vielzitierte Publikationen zu produzieren (Aksnes & Sivertsen, 2004; Bornmann, Mutz, Neuhaus, & Daniel, 2008). Hinzu kommt, dass die parametrische Statistik unter Umständen keine zuverlässigen Ergebnisse liefert, wenn der Exponent der schiefen Verteilung, aus der die Publikationen eines Untersuchungsobjekts eine „Stichprobe“ sind, kleiner drei ist (Wolfgang Glänzel, 2010; J. S. Katz, 2012). Dass parametrische Statistik dennoch anwendbar ist, wird damit begründet, dass der zentrale Grenzwertsatz eingehalten wird, wenn die Stichprobenmittelwerte normalverteilt sind. Hierfür gibt es jedoch nur eine empirische und statistisch unvollständige Untersuchung (Van Raan, 2006).

Eine Vielzahl von Funktionen ist verwendet worden, um empirische Zitationsverteilungen anzupassen (Clauset, Shalizi, & Newman, 2009; Eom & Fortunato, 2011; Laherrere & Sornette, 1998; Radicchi & Fortunato, 2008; Redner, 1998; Tsallis & de Albuquerque, 2000; Van Raan, 2001). Der Einsatz von Potenzgesetzen wird dabei favorisiert (Barabási & Albert, 1999; Peterson, Presé, & Dill, 2010; Price, 1976; Simkin & Roychowdhury, 2012).

Forschungsfragen

Im Projekt wurden statistischen Eigenschaften empirischer Zitationsverteilungen untersucht. Die Basis der Indikatoranalyse bilden vier verschiedene Modelle, die jeweils die Bedingungen für die parametrische Statistik, für die Potenzgesetz-Statistik, für beide Ansätze oder für einen der beiden Ansätze erfüllen, sowie die aus der Bibliometrie-Datenbank extrahierten zugehörigen Zitatdaten. Inhalt der Indikatoranalyse sind zwei neu entwickelte skalunenabhängige Indikatoren, die Eigenschaften der Zitationsverteilungen abbilden. Der erste der neu entwickelten Indikatoren misst Forschungsleistung anhand der Schiefe (Exponent) des Potenzgesetz-Regimes. Der zweite Indikator bewertet die Fähigkeit, weniger Publikationen zu produzieren, die wenig zitiert werden (im Vergleich mit Publikationen, die hoch zitiert werden). Diese sind mit den Indikatoren des klassischen Quadrupels verglichen worden. Neben den Indikatoren selbst bestehen weitere Forschungsfragen darin zu beantworten, bei welchem Publikationsanteil statistische Bedingungen dahingehend erfüllt sind, ob der zentrale Grenzwertsatz eingehalten wird, welche Indikatoren redundante Informationen liefern und ob die neuen Indikatoren einen Informationsmehrwert liefern.

Ergebnisse

Mit der *Scopus*-Datenbank sind die Zitationsverteilungen von 40 Ländern, 27 Disziplinen und 1.079 Land-Disziplin-Kombinationen untersucht worden. 890 (82%) der Letzteren erfüllen die Bedingungen für die parametrische Statistik (parametrisches Modell), 522 (48%) für die Potenzgesetz-Statistik (Potenzgesetz-Modell) und 1015 (94%) erfüllen eine der beiden Bedingungen (Vereinigungsmodell). Beim parametrischen Modell und beim Vereinigungsmodell wird der zentrale Grenzwertsatz eingehalten, d.h. die mittelwertbasierte Statistik ist grundsätzlich verwendbar.

Die Anzahl der Zitate, die untere Grenze für das Potenzgesetz-Verhalten und der Anteil der Publikationen im Potenzgesetz-Regime korrelieren mit der Anzahl der Publikationen, was für die Anwendung der Potenzgesetz-Statistik spricht (J. S. Katz, 2000, 2012) und die Konstruktion der vorgeschlagenen Indikatoren begründet.

Es wurde gezeigt, dass die Indikatoren des klassischen Quadrupels nicht sensitiv für statistische Anwendungsbedingungen sind. Sie korrelieren in hohem Maße über Modelle hinweg. Die skalunenabhängigen Indikatoren liefern hingegen tendenziell andere Ergebnisse, wenn die statistischen Bedingungen erfüllt sind.

Die Indikatoren des klassischen Quadrupels treffen weitgehend redundante Aussagen. Lediglich der feldnormalisierte Anteil zitierter Publikationen ist in manchen Modellen nur mittelmäßig korreliert. Insbesondere konnte gezeigt werden, dass die feldnormalisierte Zitatrate sehr stark mit dem feldnormalisierten Beitrag zu den Top 10% der meistzitierten Publikationen korreliert (Pearson-Korrelationskoeffizient mindestens 0,96 in allen Modellen). Damit wurde bestätigt, dass die feldnormalisierte Zitatrate die Fähigkeit bewertet, vielzitierte Publikationen zu produzieren.

Der skalenunabhängige Indikator liefert in Bezug auf das klassische Quadrupel nichtredundante Informationen. Die Fähigkeit, keine Publikationen zu produzieren, die wenig zitiert werden, korreliert mittel bis stark mit den Indikatoren des klassischen Quadrupels.

Implikationen für das Monitoring-Instrument

Die statistischen Anforderungen an die parametrische Statistik sind geringer als an die Potenzgesetz-Statistik. Die dadurch größere gültige Fallzahl im parametrischen Modell sowie der hohe rechentechnische und zeitliche Aufwand für die Potenzgesetz-Statistik sprechen für die parametrische Statistik. Allerdings bekräftigen die empirischen Befunde schiefer Verteilungen den Einsatz skalenunabhängiger Indikatoren.

Angesichts ihrer Redundanz scheint es nicht zielführend, sowohl die mittlere Zitatrate als auch den Beitrag zu den meistzitierten Publikationen für Anwendungsstudien parallel zu berechnen. Insgesamt plädieren wir für die Verwendung eines Indikatortripels zur Messung des Impacts:

- Anteil zitierter Publikationen,
- Mittlere Zitatrate oder Beitrag zu den meistzitierten Publikationen,
- Schiefe der Zitationsverteilung.

Für die Verwendung des Beitrags zu den meistzitierten Publikationen spricht, dass diese Indikatorfamilie frei von statistischen Anforderungen ist, da sie auf Perzentilen beruht.

2.4 Wissenschaftsdynamik

Stand der Forschung

Zur Bewertung von Wissenschaftsdynamik werden nach dem aktuellen Stand der Forschung zwei Verfahren eingesetzt. Bei der Bewertung durch visuelle Repräsentationen werden Indikatoren – meist die Anzahl der Publikationen – über den Zeitverlauf dargestellt. Um Felder oder Objekte vergleichbar zu machen, wird dabei durch den Wert des ersten Jahres geteilt. Dadurch werden Skalenunterschiede eliminiert und alle Wachstumskurven entspringen demselben Punkt. Dieses Verfahren wird in fast allen bibliometrischen Studien des KB verwendet.

Um Wachstum zu quantifizieren, wird die bewertete Wachstumsrate (Sharpe Ratio) errechnet. Diese bildet das mittlere jährliche prozentuale Wachstum ab, das zusätzlich dadurch bewertet wird, ob das Wachstum kontinuierlich oder erratisch ist (Grupp, Hinze, & Breitschopf, 2009; Schmoch, Wang, & Stoica, 2006; Sharpe, 1994). Auch die Sharpe Ratio ist bislang nur auf die Anzahl der Publikationen angewendet worden.

Skalierungsanalysen sind ein erster Schritt der Modellierung komplexer Systeme und sind in diesem Kontext auch auf Wissenschafts- und Innovationssysteme angewendet worden (Bettencourt, Kaiser, Kaur, Castillo-Chávez, & Wojick, 2008; Lane, 2009). Derartige Analysen haben die Abhängigkeit der Größe Y/N von N zum Gegenstand – beispielsweise die Abhängigkeit der Produktivität von der personellen Größe eines Objekts. Sie gehen zurück auf Katz (J. S. Katz, 2000; J. Sylvan Katz, 1999), der gezeigt hat, dass das Wissenschaftssystem skalenfrei ist – bei Veränderung der Betrachtungsgröße seine statistischen Charakteristika demzufolge beibehält. Katz hat damit gezeigt, dass die Anwendung von skalenfreien Indikatoren in wissenschaftlichen Disziplinen möglich ist.

Forschungsfragen

Im Projekt wird untersucht, ob Wachstumskurven oder -raten nicht nur auf Basis der Anzahl der Publikationen, sondern auch auf Basis weiterer Daten erstellt werden können. Diese Verfahren werden aufgearbeitet und ihre Anwendbarkeit auf andere Indikatoren demonstriert. Skalierungsanalysen erfordern hingegen tiefere Kenntnisse, um eine Interpretation der Ergebnisse zu ermöglichen. Ziel war daher, die Methode zu erschließen und Besonderheiten und Unterschiede zur Sharpe Ratio zu identifizieren. Ein weiteres Erkenntnisziel bildet die Frage, welche Wachstumsaspekte die beiden quantitativen Methoden abbilden und wie sie korrelieren.

Ergebnisse

Wachstum wurde für vier verschiedene Indikatoren abgebildet oder berechnet: Anzahl der Publikationen, kumulative Anzahl der Autoren, Anzahl der Autoren pro Publikation und Anzahl der Publikationen pro Autor. Letzterer beziffert *Produktivität*.

Zunächst trat bei der Arbeit mit *Scopus* ein Artefakt auf, das bei bibliometrischen Untersuchungen zu berücksichtigen ist: In fast allen Wissenschaftsdisziplinen nimmt die Anzahl der Zeitschriftenartikel um die Jahre 2002 bis 2003 ab, obwohl monotonen Wachstum erwartet wird und auf derartiger Aggregationsebene nicht mit solch heftigen Fluktuationen zu rechnen ist. Elsevier nahm zu den Ursachen auf Anfrage nicht Stellung, wir vermuten jedoch einen Zusammenhang zur Veränderung der Klassifizierungspolitik in diesem Zeitraum. Dieses Artefakt ist ein zwingender Grund für die Berücksichtigung des Datenbankwachstums bei der Berechnung des Publikationswachstums.

Eine Bewertung der Dynamik auf Basis von visuellen Repräsentationen und Indikatorik ist empfehlenswert. Wachstumskurven geben einen Eindruck von den Daten und weisen auf Besonderheiten wie unterschiedliche Wachstumsregime hin, die dann zusätzlich mit Hilfe der Indikatoren differenziert quantitativ untersucht werden sollten. Normalisierte Wachstumskurven ermöglichen Vergleiche von Feldern oder Objekten, tendieren jedoch auch dazu, die Unterscheidung von Details durch Überlagerung von Kurven zu erschweren. Die Abbildung von Wachstum wurde für verschiedene Indikatoren demonstriert.

Die Skalierungsanalyse ermöglicht eine grundlegend andere Perspektive auf Wissenschaftswachstum. Beziffert wird das gleichbleibende Systemverhalten *über* mehrere Größen hinweg, nicht das mittlere Wachstum *von* verschiedenen Größen. Bei der Skalierungsanalyse ist der Zeitfaktor indirekt in der Größe der Variable N enthalten. Positives Wachstum drückt sich in einem Skalenerunterschied aus, umgekehrt jedoch nicht zwingend. Da der Skalierungsexponent in mathematischer Beziehung zu Änderungen der jährlichen Wachstumsrate steht, auf der die Sharpe Ratio beruht, sind beide Wachstumsindikatoren dennoch korreliert (Pearson Korrelationskoeffizient $\geq 0,67$).

Implikationen für das Monitoring-Instrument

Sofern *Scopus* verwendet wird, sollte dieses erst ab dem Publikationsjahrgang 2004 geschehen, um die Diskontinuität von 2002-2003 zu umgehen. Verwendet werden sollten außerdem nicht nur der Publikationstyp „Article“, sondern ebenfalls Reviews und Beiträge zu Konferenzbänden.

Während die Sharpe Ratio für sämtliche Indikatoren berechnet werden kann, eignen sich Skalierungsanalysen nur für bivariate Indikatoren wie etwa Produktivität (Anzahl Publikationen pro Autor) oder Impact (Anzahl Zitate pro Publikation).

Eine Korrektur für das Datenbankwachstum findet bei der bewerteten Wachstumsrate per Definition statt. Bei der Skalierungsanalyse ist solch eine Korrektur möglich, indem der Wachstumsindikator (Skalierungsexponent) des zu quantifizierenden Objekts durch den der Datenbank geteilt wird. Die Referenz sollte jedoch mit Bedacht gewählt werden. Beispielsweise sollte darauf geachtet werden, dass das Objektwachstum nicht mit disziplinfremdem Datenbankwachstum normalisiert wird, da *Scopus* in den Disziplinen unterschiedlich wächst.

Die Ergebnisse verdeutlichen, dass die zunehmende Bedeutung von Teams bei der Wissensproduktion zumindest auf der Ebene von Disziplinen ein Produktivitätswachstum verhindert. Eine mögliche Erklärung dafür könnte darin liegen, dass Skalierungsexponenten zur Quantifizierung von Wachstum besser für emergente Felder geeignet sind.

3 Literatur

- Aksnes, D. W., & Sivertsen, G. (2004). The effect of highly cited papers on national citation indicators. *Scientometrics*, 59(2), 213–224.
- Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *science*, 286(5439), 509–512.
- Bettencourt, L. M., Kaiser, D. I., Kaur, J., Castillo-Chávez, C., & Wojick, D. E. (2008). Population modeling of the emergence and development of scientific fields. *Scientometrics*, 75(3), 495–518.
- Bornmann, L. (2013). How to analyze percentile citation impact data meaningfully in bibliometrics: The statistical analysis of distributions, percentile rank classes, and top-cited papers. *Journal of the American Society for Information Science and Technology*.
- Bornmann, L., Mutz, R., Neuhaus, C., & Daniel, H.-D. (2008). Citation counts for research evaluation: standards of good practice for analyzing bibliometric data and presenting and interpreting results. *Ethics in Science and Environmental Politics*, 8(1), 93–102.
- Chinchilla-Rodríguez, Z., Vargas-Quesada, B., Hassan-Montero, Y., González-Molina, A., & Moya-Anegón, F. (2010). New approach to the visualization of international scientific collaboration. *Information Visualization*, 9(4), 277–287.
- Clauset, A., Shalizi, C. R., & Newman, M. E. (2009). Power-law distributions in empirical data. *SIAM review*, 51(4), 661–703.
- Ding, Y., Foo, S., & Chowdhury, G. (1998). A bibliometric analysis of collaboration in the field of information retrieval. *International Information and Library Review*, 30, 367–376.
- Eom, Y.-H., & Fortunato, S. (2011). Characterizing and modeling citation dynamics. *PloS one*, 6(9), e24926.
- Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics*, 51(1), 69–115.
- Glänzel, W. (2002). Coauthorship patterns and trends in the sciences (1980-1998): A bibliometric study with implications for database indexing and search strategies. Katholieke Universiteit Leuven. Abgerufen von <http://ideas.repec.org/p/ner/leuven/urnhdl123456789-98596.html>
- Glänzel, W., & Schubert, A. (2001). Double effort = Double impact? A critical view at international co-authorship in chemistry. *Scientometrics*, 50(2), 199–214. doi:10.1023/A:1010561321723
- Glänzel, W., Schubert, A., & Czerwon, H. J. (1999). A bibliometric analysis of international scientific cooperation of the European Union (1985–1995). *Scientometrics*, 45(2), 185–202.
- Glänzel, Wolfgang. (2010). On reliability and robustness of scientometrics indicators based on stochastic models. An evidence-based opinion paper. *Journal of Informetrics*, 4(3), 313–319.
- Glänzel, Wolfgang, & Thijs, B. (2011). Using ‘core documents’ for the representation of clusters and topics. *Scientometrics*, 88, 297–309.
- Grupp, H., Hinze, S., & Breitschopf, B. (2009). Defining regional research priorities: a new approach. *Science and Public Policy*, 36(7), 549–559.
- Guerrero Bote, V. P., Olmeda-Gómez, C., & de Moya-Anegón, F. (2013). Quantifying the benefits of international scientific collaboration. *Journal of the American Society for Information Science and Technology*, 64(2), 392–404. doi:10.1002/asi.22754
- Hsu, J., & Huang, D. (2011). Correlation between impact and collaboration. *Scientometrics*, 86(2), 317–324.
- Katz, J. S. (2000). Scale-independent indicators and research evaluation. *Science and Public Policy*, 27(1), 23–36.

- Katz, J. S. (2012). *Scale-Independent Measures: Theory and Practice*.
- Katz, J. S., & Hicks, D. (1997). How much is a collaboration worth? A calibrated bibliometric model. *Scientometrics*, 40(3), 541–554.
- Katz, J. Sylvan. (1999). The self-similar science system. *Research policy*, 28(5), 501–517.
- Laherrere, J., & Sornette, D. (1998). Stretched exponential distributions in nature and economy: “fat tails” with characteristic scales. *The European Physical Journal B-Condensed Matter and Complex Systems*, 2(4), 525–539.
- Lancho-Barrantes, B. S., Guerrero-Bote, V. P., & de Moya-Anegón, F. (2012). Citation increments between collaborating countries. *Scientometrics*, 1–15.
- Lane, D. A. (2009). *Complexity perspectives in innovation and social change* (Bd. 7). Springer.
- Lewinson, G., & Cunningham, P. (1991). Bibliometric studies for the evaluation of trans-domestic research. *Scientometrics*, 21(2), 223–244.
- Lundberg, J. (2007). Lifting the crown: Citation z-score. *Journal of Informetrics*, 145–154.
- Luukkonen, T., Tijssen, R. J. W., Persson, O., & Sivertsen, G. (1993). The measurement of international scientific collaboration. *Scientometrics*, 28(1), 15–36.
- Maghrebi, M., Abbasi, A., Amiri, S., Monsefi, R., & Harati, A. (2011). A collective and abridged lexical query for delineation of nanotechnology publications. *Scientometrics*, 86(1), 15–25.
- Mogoutov, A., & Kahane, B. (2007). Data search strategy for science and technology emergence: A scalable and evolutionary query for nanotechnology tracking. *Research Policy*, 36, 893–903.
- Narin, F., Stevens, K., & Whitlow, E. S. (1991). Scientific co-operation in Europe and the citation of multinationally authored papers. *Scientometrics*, 21(3), 313–323.
- Noyens, E., Buter, R., van Raan, A., Schmoch, U., Heinze, T., Hinze, S., & Rangnow, R. (2003). *Mapping Excellence in Science and Technology across Europe - Life Sciences*. Report of project EC-PPLS CT-2002-0001, European Commission.
- Pao, M. L. (1981). Co-authorship as communication measure. *Library Research*, 2(4), 327–338.
- Persson, O., Glänzel, W., & Danell, R. (2004). Inflationary bibliometric values: The role of scientific collaboration and the need for relative indicators in evaluative studies. *Scientometrics*, 60(3), 421–432.
- Peterson, G. J., Pressé, S., & Dill, K. A. (2010). Nonuniversal power law scaling in the probability distribution of scientific citations. *Proceedings of the National Academy of Sciences*, 107(37), 16023–16027.
- Price, D. de S. (1976). A general theory of bibliometric and other cumulative advantage processes. *Journal of the American Society for Information Science*, 27(5), 292–306.
- Radicchi, F., & Castellano, C. (2011). Rescaling citations of publications in physics. *Physical Review E*, 83.
- Radicchi, F., & Fortunato, S. (2008). Universality of citation distributions: Toward an objective measure of scientific impact.
- Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B-Condensed Matter and Complex Systems*, 4(2), 131–134.
- Schmoch, U., Wang, J., & Stoica, R. (2006). *Research and Development in the Turkish Landscape of Science*.
- Schubert, A., & Braun, T. (1986). Relative indicators and relational charts for comparative assessment of publication output and citation impact. *Scientometrics*, 9.

- Sharpe, W. F. (1994). The sharpe ratio. *Journal of portfolio management*, 21, 49–49.
- Simkin, M. V., & Roychowdhury, V. P. (2012). Theory of citing. In *Handbook of Optimization in Complex Networks* (S. 463–505). Springer.
- Singh, J. (2005). Collaborative networks as determinants of knowledge diffusion patterns. *Management science*, 51(5), 756–770.
- Smolinsky, L., & Lercher, A. (2012). Citation rates in mathematics: a study of variation by sub-discipline. *Scientometrics*, 91, 911–924.
- Strotmann, A., & Zhao, D. (2010). Combining commercial citation indexes and open-access bibliographic databases to delimit highly interdisciplinary research fields for citation analysis. *Journal of Informetrics*, 4, 194–200.
- Strotmann, A., Zhao, D., & Bubela, T. (2009). A multi-database approach to field delineation. *Proceedings of ISSI 2009*, 631–635.
- Tijssen, R. J., Visser, M. S., & Van Leeuwen, T. N. (2002). Benchmarking international scientific excellence: Are highly cited research papers an appropriate frame of reference? *Scientometrics*, 54(3), 381–397.
- Tsallis, C., & de Albuquerque, M. P. (2000). Are citations of scientific papers a case of nonextensivity? *The European Physical Journal B-Condensed Matter and Complex Systems*, 13(4), 777–780.
- Van Raan, A. F. (2001). Competition amongst scientists for publication status: Toward a model of scientific publication and citation distributions. *Scientometrics*, 51(1), 347–357.
- Van Raan, A. F. (2006). Statistical properties of bibliometric indicators: Research group indicator distributions and correlations. *Journal of the American Society for Information Science and Technology*, 57, 408–430.
- Waltman, L., van Eck, N. J., van Leeuwen, T. N., Visser, M. S., & van Raan, A. F. (2011). Towards a new crown indicator: Some theoretical considerations. *Journal of Informetrics*, 5(1), 37–47.
- Zhao, Q., & Guan, J. (2011). International collaboration of three ‘giants’ with the G7 countries in emerging nanobiopharmaceuticals. *Scientometrics*, 87(1), 159–170.
- Zitt, M., & Bassecoulard, E. (o. J.). Delineating complex scientific fields by an hybrid lexical-citation method: An application to nanosciences. *Information Processing & Management*, 42, 1513–1531.